

Methode zur Gestaltung sicherer präsriptiver Systeme für die Therapieunterstützung

Philipp Skowron

Institut für Wirtschaftsinformatik
Universität Leipzig
Leipzig, Deutschland
skowron@wifa.uni-leipzig.de

Abstract—Die zunehmende Geschwindigkeit der Digitalisierung im Gesundheitsbereich bietet diesem neue Möglichkeiten, stellt die Branche aber auch vor neue Herausforderungen. Mithilfe von präsriptiven Systemen können Langzeittherapien unterstützt werden, um deren Therapietreue zu verbessern. Hierfür müssen jedoch bei der Konzeption und Entwicklung solcher Systeme verschiedenste Faktoren, wie die Therapietreue und die Algorithmensicherheit, mit einbezogen werden, um einen gesamtgesellschaftlichen Mehrwert schaffen zu können. Die Arbeit stellt eine Methode zur Konzeption und Entwicklung therapieunterstützender präsriptiver Systeme vor, wobei sowohl die Therapietreue als auch die Sicherheit der Entscheidungen im Vordergrund steht.

Keywords—Therapieunterstützung; Präsriptive Systeme; sichere Entscheidungsunterstützung

I. EINLEITUNG

Digitalisierung und Medizin waren schon immer eng miteinander verbunden gewesen. Behandlungsmöglichkeiten und -qualität steigen mit den technischen Möglichkeiten an, die durch die Digitalisierung geboten werden. Dies zeigt sich u. a. in den Marktpotentialen für digitale Lösungen im Medizinbereich. Gerade die Bereiche Wireless Health und mHealth sollen mit zusammen 156 Mrd. USD Marktanteil (im Jahr 2020) das Monitoring und die medizinische und therapeutische Begleitung von Patienten zukünftig vereinfachen [1]-[2]. Technische Lösungen aus dem IoT-Sektor für mobile Anwendungen können mit Sensoren und Aktuatoren ausgestattet werden, die das Umfeld und den Patienten überwachen.

Während der dadurch umfänglicheren medizinischen Begleitung können Patienten mit chronischen Erkrankungen besser angeleitet und deren Therapietreue, also deren Einhaltung aller therapeutischen Maßnahmen, gestärkt werden, was grundsätzlich zu einer verbesserten Gesundheitsversorgung führt. Hier bieten sich vor allem therapeutische Unterstützungssysteme an, die mithilfe von entscheidungsunterstützenden Funktionalitäten dem Patienten zur Seite stehen, auch wenn das pflegende medizinische Personal nicht direkt vor Ort ist [3]. Dies adressiert zugleich den Trend hin zur alternden Gesellschaft, welcher durch die stagnierende Anzahl an qualifiziertem Pflegepersonal, einen Mangel einer halben Million Pflegefachkräften, zusätzlich verstärkt wird [4]. Die hohe Beanspruchung des vorhandenen Pflegepersonals in ambulanten und stationären Bereichen führt zu einer körperlichen, zeitlichen und psychischen Überlastung der

Pflegenden [5]. Dabei senkt die immense Auslastung des Personals die Qualität der Betreuung und Therapie der Patienten.

Eingangs erwähnte Sensor- und Aktorsysteme können zur Entlastung der Pflegekräfte während ihrer täglichen Aufgaben beitragen [6]. Das hohe Marktpotential in der Gesundheitsbranche für sensorbasierte Assistenzsysteme im Therapiebereich [3] treibt die Entwicklung dieser massiv voran. Jedoch sind Standards und allgemein anerkannte Methoden zur Entwicklung therapieunterstützender Entscheidungsunterstützungssysteme, aufgrund der differenzierten Herangehensweise seitens der Entwicklung und Sicherstellung von Funktionalität und Nutzen, kaum zu erkennen [7]-[8].

Eine explizite Betrachtung der Sicherheit der Entscheidungen und Therapieunterstützung wird in bereits etablierten Prozessen und Methoden nur ungenügend abgebildet. Hieraus ergibt sich der Bedarf zur Entwicklung einer Methode für die Entwicklung von entscheidungsunterstützenden Therapieunterstützungssystemen, welche sowohl therapeutische und Personen-individuelle Charakteristika, als auch die explizite Betrachtung der Sicherheit von angewendeten präsriptiven Verfahren in den Prozess einbezieht.

Zur wissenschaftlichen und praktischen Einordnung von entscheidungsunterstützenden Therapiesystemen kann eine Klassifizierung der World Health Organization (WHO) für digitale Technologien im Gesundheitswesen zu Rate gezogen werden [9]. Hierin werden digitale Interventionen zur Lösung oder Milderung bestehender Herausforderungen im Gesundheitswesen definiert. Dabei werden die Interventionen in vier verschiedene Kategorien unterteilt, welche verschiedene Rollenträger und Technologien im Gesundheitssystem ansprechen. Hierunter fallen die Patienten des Gesundheitssystems, die Krankenversicherungen, die Gesundheitsmanager im System und die Technologien zur Datensammlung, -verwendung und zum Datenaustausch. Zusätzlich definiert die WHO bestehende Herausforderungen im Gesundheitssystem sowie Systemkategorien, in die eine weitere Unterteilung der Lösungen stattfinden kann.

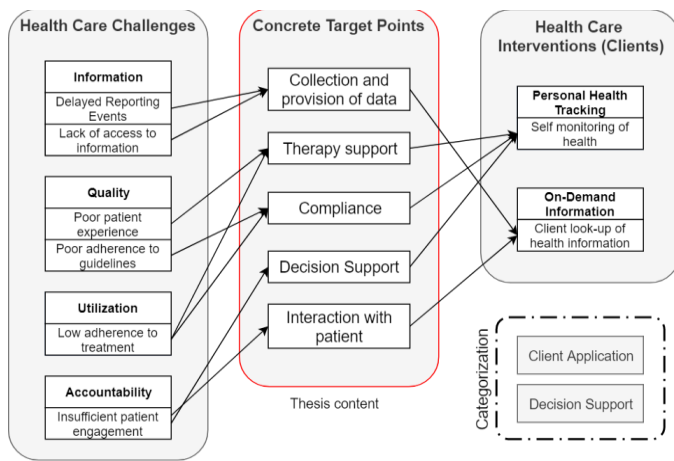


Fig. 1. Einordnung Dissertation in Kontext der WHO

Wie in Fig. 1 dargestellt, teilt sich die Einordnung in drei Hauptteile. Der Inhalt und das Ziel der Arbeit rot umrandet dargestellt. Lesbar ist die Einordnung wie folgt: „Durch die Sammlung und Bereitstellung von Daten über den Patienten (Target Points), können die Herausforderungen der zu späten Mitteilung von Ereignissen und der generelle Mangel an Informationen über die Patienten adressiert werden (Health Care Challenges). Dabei lässt sich die Sammlung und Bereitstellung der Daten in die Gruppe der patientenbezogenen digitalen Interventionen einsortieren, die geprägt ist durch personalisiertes Monitoring mithilfe von körpernaher Sensorik.“ Die Kategorisierung ist zusätzlich eine von der WHO vorgeschlagene Zuordnung der Interventionen zu Kategorien von Informations- und Telekommunikations-Systemen und bietet eine weitere Unterteilung der Systeme auf technischer Seite [9].

Ersichtlich ist in der Zuordnung zu den Health Care Interventions, dass ausschließlich patientenbezogene Interventionen betrachtet werden. Somit ist die zu entwickelnde Methode darauf ausgelegt, Systeme zu entwickeln, die für die Nutzung durch den Patienten konzipiert werden.

II. FORSCHUNGSFRAGEN

Zur Entwicklung einer solchen benannten Methode, die die angesprochenen Probleme hinreichend entgegen kann, ist der Rahmen einer übergeordneten Forschungsfrage wie folgt zu wählen: **Wie muss eine Methode gestaltet sein, um ein Anwendungssystem für die therapeutische Unterstützung zu entwickeln, welches die Sicherheit der Entscheidungen bei der Therapieunterstützung explizit unterstützt?**

Wie aus der übergeordneten Fragestellung hervorgeht, soll eine Methode entwickelt werden, die eine explizite Betrachtung von Algorithmensicherheit in Systemen in der Therapieunterstützung zulässt. Dabei liegt der Fokus auf der Konzeption und Umsetzung solcher Unterstützungssysteme. Um die Methode entwickeln und den Rahmen konkreter betrachten zu können, ist die Definition von drei Unterforschungsfragen nötig, die wie folgt lauten:

- **RQ1:** Welche Ansätze für die sichere Gestaltung von entscheidungsunterstützenden Therapieunterstützungssystemen existieren bereits?

- **RQ2:** Wie müssen die Anforderungen an eine Methode gestaltet sein, damit diese sichere Entscheidungsfindungen bei therapeutischen Unterstützungssystemen gewährleisten?
- **RQ3:** Welche Optimierungen lassen sich für die Sicherheit in therapeutischen Unterstützungssystemen erzielen?

Wie vor allem die RQ3 beschreibt, liegt vor allem der Fokus der Methode, neben dem Einbezug von Therapieelementen und der Adhärenz, auf der Sicherstellung von nicht schadhaften, transparenten Entscheidungen, die von den zu entwickelnden Systemen, je nach Zielstellung, autonom getroffen werden können. Da eine Therapieunterstützung immer direkte Auswirkungen auf Menschen hat, die therapiert werden sollen, ist Sicherheit der Entscheidungen ein wichtiger Schwerpunkt bei solchen Entwicklungen.

III. FORSCHUNGSRAHMEN, METHODIK & ARTEFAKTE

Der Forschungsrahmen wurde nach Österle et al. erstellt und liegt dem Ansatz der gestaltungsorientierten Wirtschaftsinformatik zu Grunde. Entsprechend des Ansatzes werden die Phasen der Analyse, des Entwurfs, der Evaluation und der Diffusion durchlaufen. In der Bearbeitung der Phasen werden verschiedene Arten von Artefakten geschaffen, die Methoden, Modelle, Konstrukte oder Instanzen sein können. Diese unterliegen den Prinzipien der Abstraktion, Originalität, Begründung und des Nutzens [10].

Zur Bearbeitung der Forschungsfragen wurde zu Beginn eine systematische Literaturanalyse nach dem Schema von vom Brocke et al. [12] erstellt, die sowohl Herausforderungen und Probleme bei der Erstellung von präskriptiven Systemen identifizierte als auch bestehende Ansätze und Methoden aufdeckte. Alle gesammelten Erkenntnisse aus der Literaturanalyse wurden durch Gespräche mit Experten [13][14] erweitert und vertieft, um sowohl wissenschaftliche als auch praktische Ansichten in die Bearbeitung der Anforderungen mit einfließen zu lassen.

Sowohl die systematische Literaturanalyse als auch die Experteninterviews dienen der Klärung der Fragestellung aus RQ1.

Auf Basis der Probleme, Herausforderungen, Methoden und Ansätze aus der systematischen Literaturanalyse konnte ein für die Entwicklung von entscheidungsunterstützenden Therapieunterstützungssystemen benötigter Anforderungskatalog realisiert werden. Dieser bildet alle Anforderungen ab, gegliedert in drei verschiedene Anforderungsbereiche, die für eine Methode von Bedeutung sind, um die untergeordnete Forschungsfrage RQ2 beantworten zu können.

Methodisch wurde der Anforderungskatalog unter Zuhilfenahme des Requirements Engineerings [15] erstellt.

Bei der Erstellung des Anforderungskatalogs wurden drei Anforderungsbereiche definiert. In den allgemeinen Anforderungen an Methoden werden jene Anforderungen eingeordnet, die eine Methode unabhängig von der Domänenspezifität besitzen muss. Diese sollen die Unabhängigkeit und Flexibilität der Methode gewährleisten, um einen Grad der Generalisierung zur Erfüllung des Prinzips der Abstraktion zu erreichen.

Ein weiterer Bereich in den Anforderungen ist jener für therapieunterstützende Anwendungen. Alle hier eingeordneten Anforderungen dienen der spezifischen Ausrichtung der Methode in Richtung von therapieunterstützenden Systemen und somit der Abdeckung aller therapeutisch-relevanten Anforderungen wie bspw. Therapietreue und die therapeutische Ausrichtung.

Der dritte Bereich der Anforderungen definiert die Anforderungen an entscheidungsunterstützende Algorithmen. Unter anderem wird hier explizit die Sicherheit solcher Algorithmen betrachtet, um einen sicheren Umgang mit jenen auch in voll-autonomen Einsatzgebieten gewährleisten zu können und somit auch die Sicherheit der Nutzer zu garantieren.

Auf Basis dieses Anforderungskatalogs wurden die in der systematischen Literaturanalyse identifizierten Methoden und Ansätze auf deren Vollständigkeit gegenüber den Anforderungen geprüft [12]. Es ergab sich aus dem Vergleich der Anforderungen mit den Methoden und Ansätzen, dass keiner der Ansätze ein vollständiges Abbild der Anforderungen zeichnen konnte (siehe Fig. 2). Die Ansätze werden im späteren Verlauf noch einmal näher benannt.

	An1	An2	An3	An4	An5	An6	An7	An8
A1	+	+	+	+	+	+	+	+
A2	+	+	+	+	+	+	+	+
A3	+	+	+	+	+	+	+	+
A4	+	+	+	+	+	+	+	+
A5	+	+	+	+	+	+	+	+
T1	+	+	+	+	+	+	+	+
T2	+	+	+	+	+	+	+	+
T3	+	+	+	+	+	+	+	+
P1	+	+	+	+	+	+	+	+
P2	+	+	+	+	+	+	+	+
P3	+	+	+	+	+	+	+	+
P4	+	+	+	+	+	+	+	+
P5	+	+	+	+	+	+	+	+
P6	+	+	+	+	+	+	+	+

Fig. 2. Ergebnis Anforderungsauswertung (AnX – Ansätze & Methoden; Ax, Tx, Px – Anforderungen der einzelnen Gruppen)

Demnach ist die Entwicklung einer Methode zur Erfüllung des definierten Anforderungskatalogs nötig.

Zur Erstellung der in Fig. 3 dargestellten Methode wurde das Method Engineering [16] als methodischer Rahmen genutzt.

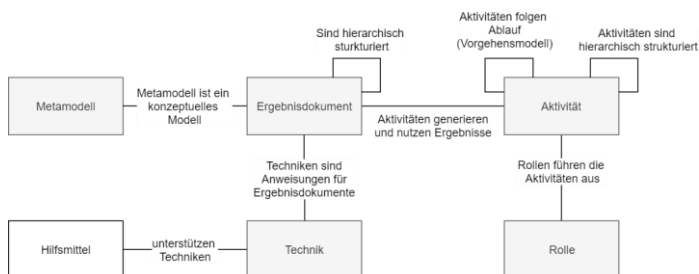


Fig. 3. Elemente einer Methode nach dem Method Engineering

Die Summe aus verschiedenen Modellen definiert in ihrer Gesamtheit die Methode im Ganzen. Dabei bildet das Vorgehensmodell, welches aus Aktivitäten, Ergebnissen und Rollen besteht, ein zentrales Artefakt der Methode. Weiterhin definiert ein Technikmodell die anzuwendenden Techniken bei der

Generierung der Ergebnisse. Das Metamodell der Methode ist ein konzeptuelles Modell, welches die Gestaltungsobjekte für die Erstellung eines Modells beschreibt und diese in Beziehung zueinander setzt.

Nach der vollständigen Erfüllung aller Anforderungen in der Methode wird eine Evaluation im Rahmen eines Forschungsprojektes anhand eines Anwendungsfalles durchgeführt. Dies dient der Erfüllung des Ablaufes im Rahmen der gestaltungsorientierten Wirtschaftsinformatik einerseits und andererseits der Beantwortung der RQ3.

IV. PROBLEMANALYSE

Für die Entwicklung eines solchen entscheidungsunterstützenden Therapieunterstützungssystems fallen verschiedene Herausforderungen an, die es zu bewältigen gilt und welche im Folgenden kurz dargestellt werden.

A. Herausforderungen an eine Methode

Da die Therapieunterstützung darauf abzielt, Therapien bei chronischen Erkrankungen in ihrer Wirksamkeit zu unterstützen, spielt die Betrachtung der Adhärenz eine entscheidende Rolle. Weiterhin ist es wichtig, die Sicherheit der Entscheidungsunterstützung mit einzubeziehen, um die Sicherheit der Nutzer und Patienten während der Nutzung des Systems gewährleisten zu können. Damit die Adhärenz aufgenommen und die Entscheidungsunterstützung überhaupt durchgeführt werden kann, muss der Datenschutz als eine Herausforderung mit einbezogen werden.

Folgend sollen die drei Herausforderungen kurz näher erläutert werden.

Adhärenz oder zu Deutsch Therapietreue, ist definiert als der Grad der Genauigkeit, wie ein Patient die Anweisungen des medizinischen Personals Folge leistet. Dabei hat sich das Bild der Therapietreue über die Jahre hin geändert, von einer autoritären Weisung des Mediziners an den Patienten, hin zu einem kooperativen Gestalten von Therapieansätzen. Vor allem bei chronischen Erkrankungen ist eine langfristige Einhaltung von therapeutischen Maßnahmen grundsätzlich schwierig. Nach einer Studie über die Adhärenz der WHO werden gerade einmal 50% aller Therapien genau eingehalten, was Schluss endlich als ein Therapieerfolg gewertet werden kann. Die anderen 50% hingegen gelten als therapeutischer Misserfolg, der z. B. in stationären Krankenhausaufenthalten endet. Dadurch entstehen jährlich weltweit in Nicht-Entwicklungsländern Kosten von bis zu 300 Mrd. US \$, 10 Mrd. € jährlich sind es in Deutschland [17]-[18].

Um diese Zahl zu reduzieren, kann durch geeignete Messungen und gezielte Maßnahmen eine langfristige Therapietreue erreicht werden. Jedoch kann es hier keine allgemeine Regelung geben, da jede Situation bei jedem Patienten auf individuellen Rahmenbedingungen fußt. Hierfür hat die WHO einen Kriterienkatalog herausgegeben, der die Kategorien für die Erfassung der individuellen Umstände festlegt und definiert [17], [19].

Wie aus Fig. 4 ersichtlich, lassen sich die fünf Gruppen, definiert durch die WHO, in drei Obergruppen zusammenfassen. Diese betrachten patientenbezogene Faktoren, wie bspw. das Alter des Patienten, was ein entscheidender demographischer und medizinischer Fakt ist, Faktoren aus der Erkrankung und

Therapie, wie die Dauer der Therapie oder die Schwere der Erkrankung, sowie Faktoren, die von der Betreuung des Arztes und den Rahmenbedingungen des Gesundheitssystems abhängig sind.

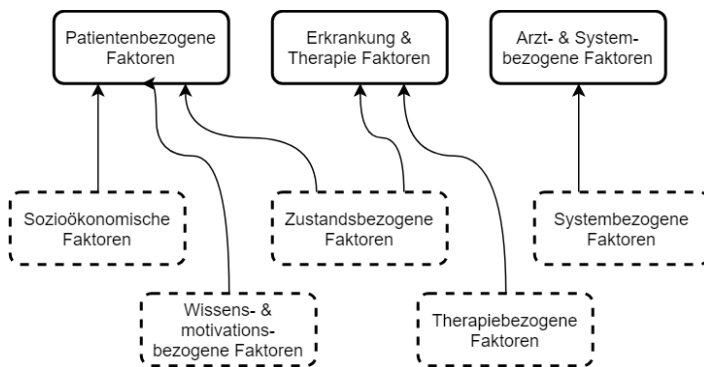


Fig. 4. Adhärenzfaktoren Gruppen (in Anlehnung an [17] und [20])

Fig. 5 gibt einen Einblick in den Verlauf von Erkrankungs- und Therapiefaktoren, die die Therapietreue beeinflussen. So sind vor allem die Dauer der Therapie, die Anzahl der Applikationen pro Tag und der Schweregrad der Erkrankung entscheidend, wie groß die Wahrscheinlichkeit für eine schlechte Therapietreue ist.

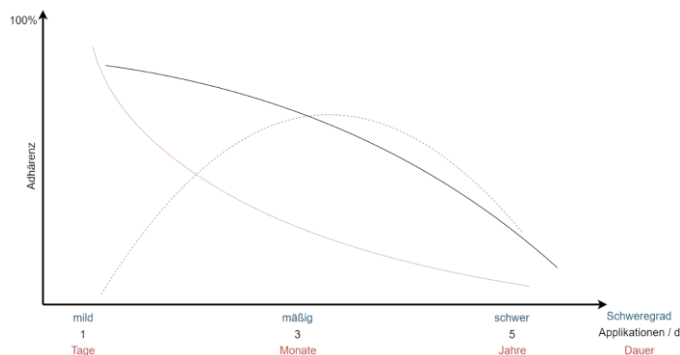


Fig. 5. Adhärenz Verlauf unter verschiedenen Faktoren (in Anlehnung an [17])

Wurde die Therapietreue aufgenommen und die Zielgruppe bewertet, kann theoretisch betrachtet ein entsprechendes System konzeptioniert werden, welches bestimmte Umstände der Adhärenz adressiert, z. B. den Mangel an Bewegung, um eine Therapietreue langfristig zu unterstützen. Ganz gleich, wie dieses System aussieht, sobald autonome oder teilautonome Entscheidungsfindung genutzt wird, ist die explizite Betrachtung der Sicherheit für den Patienten / Nutzer während der Nutzung ein entscheidendes Kriterium.

Um eine ganzheitliche Sicherheitsbetrachtung durchzuführen, müssen sowohl die prozessuale Sicherheit als auch die Sicherheit vor unberechtigten Zugriff und Veränderung betrachtet werden [21]-[22]. Weiterführend soll auf die prozessuale Sicherheit des Algorithmus, also die Sicherheit der Entscheidungsunterstützung, näher eingegangen werden. Aspekte der Sicherheit vor unberechtigten Zugriff sind allgemeine

Datensicherheitsfragen, wie Verschlüsselung der Daten, und sind in der vorliegenden Arbeit als selbstverständlich zu betrachten.

Algorithmen, welche Entscheidungen auf Basis von Daten treffen, um ein bestimmtes Ziel im optimalen Grad zu erreichen, bezeichnet man als präskriptive Algorithmen. Präskriptive Algorithmen ordnen sich gegenüber vorhersagenden oder auswertenden Algorithmentypen als separater Typ ein. Sie verhalten sich autonomer als die anderen beiden (benötigen weniger Expertenwissen) und betrachten die Zukunft, sind aber in ihren Freiheitsgraden sehr komplex und, vor allem unter Betrachtung aktuellen maschinellen Lernens wie neuronale Netze, nicht oder nur schwer selbsterklärend in ihrer Entscheidungsfindung [23]-[24].

Da zunehmend in den vergangenen Jahren präskriptive Algorithmen, wie das bestärkende Lernen, forciert wurden, wurden auch eine Reihe von Sicherheitsproblemen identifiziert, die eine Beachtung bei der Entwicklung eines präskriptiven Systems benötigen.

Beispielhaft kann der „distributional shift“ genannt werden, welcher bei Anwendungen in der realen Welt grundsätzlich vorhanden ist. Hierbei wird ein System auf Basis von Daten und Informationen erstellt, welche nur ein Abbild der realen Welt darstellen (bspw. Verkehrsflüsse an Kreuzungen). Optimiert die präskriptive Anwendung in diesem Abbild seine Zielfunktion ausreichend, kann diese in der realen Welt genutzt werden. Ändern sich jedoch Rahmenbedingungen (z. B. Verkehrsstärke, neue Kreuzungen oder neue Fahrzeugtypen), die die Anwendung noch nie gesehen hat, kann es zu unvorhergesehenen und ggf. sicherheitskritischen Verhalten kommen [25]. Zu diesem Beispiel kommen noch acht weitere hinzu, welche betrachtet und behandelt werden müssen, um ein sicheres präskriptives System erstellen zu können.

Als letzte große Herausforderung bei der Erstellung von präskriptiven Therapieunterstützungssystemen ist der Datenschutz und die Einhaltung dieses zu erwähnen.

Fallen Daten an, welche eine Person direkt oder indirekt identifizieren können, greift die Datenschutzgrundverordnung (DSGVO). Gerade im medizinischen Bereich, auch wenn man selbst kein Medizinprodukt nach der Medical Device Regulation (MDR) erstellt, sind medizinische Daten hochgradig datenschutzrelevant. Ist mit der Erfassung dieser Daten auch eine Verarbeitung dieser verbunden (z. B. die Auswertung durch maschinelles Lernen), ist fast immer eine Datenschutzfolgenabschätzung durchzuführen, welche die Risiken für die schützenswerten Daten der Personen identifiziert, bewertet und Maßnahmen zur Linderung der Risiken definiert (Art. 35 Abs. 1 DSGVO).

Zu den datenschutzrechtlichen Aspekten ist eine Abgrenzung zu der MDR durchzuführen. Ist dies nicht direkt möglich, weil bspw. die MDR aufgrund der Funktionalität des zu entwickelnden Systems greift, muss eine Klassifizierung des Systems nach der MDR durchgeführt werden [26]. Dies kann im ungünstigsten Fall dazu führen, dass das System als Medizinprodukt eingestuft wird und zusätzlich zu den Regelungen der DSGVO, weitere weitreichende Dokumentations- und Studienpflichten auferlegt bekommt.

B. Anforderungen an eine Methode

Nachdem die Herausforderungen erklärt wurden, werden im Folgenden die Anforderungen an ein solches System beschrieben, die die Herausforderungen mit einbeziehen.

Wie bereits eingangs beschrieben, wurden die Anforderungen im Rahmen der *RQ1* in einer systematischen Literaturanalyse identifiziert und für die Beantwortung der *RQ2* beschrieben und in einem Anforderungskatalog zusammengefasst. Dabei gliedern sich die Anforderungen in drei Bereiche, welche verschiedene Aspekte der zu entwickelnden Methode definieren.

Der allgemeine Teil der Anforderungen soll die Unabhängigkeit der Domänenspezifität einer Methode gewährleisten sowie Charakterzüge fordern, die in allen Methoden enthalten sind.

Die Anforderung der **iterativen Fähigkeit** (A1) des Ansatzes bedingt, dass der Ansatz in Teilergebnisse zerlegt werden kann und Rücksprünge zu früheren Phasen des Prozesses erlauben soll, um frühzeitig auf Nutzer-Feedback und auf veränderte Anforderungen eingehen zu können [27]. Die **Flexibilitätsanforderung** (A2) beschreibt, dass ein Ansatz generisch sein muss und nicht an bestimmte Anwendungsfälle oder Technologien geknüpft sein darf [16]. Zudem ist die **Nachvollziehbarkeit** (A3), also die Rückverfolgbarkeit von Funktionalitäten zu Anforderungen eine notwendige Anforderung, um auch Sicherheit im System gewährleisten zu können [28]. Die Zielerreichung und damit verbunden die definierten Ziele in dem Ansatz, müssen operationalisierbar sein, um deren Fertigstellungs- und Erfolgsgrad messen zu können. Dies schreibt die Anforderung der **Messbarkeit** (A4) vor und ist vor allem in Verbindung mit dem iterativen Design des Ansatzes aus der ersten Anforderung zu sehen [16]. Die letzte Anforderung aus dem allgemeinen Teil des Anforderungskatalogs bezieht sich auf die **Ergebnisorientiertheit** (A5) eines Ansatzes. Sie stellt sicher, dass alle Aktivitäten konkrete und im Prozess relevante Artefakte erstellen und somit einen Mehrwert zur Erreichung des Zieles bieten [16].

Im zweiten Teil des Anforderungskatalogs finden sich die Anforderungen, die im Zusammenhang mit der Erstellung therapieunterstützender Systeme stehen und somit einen therapeutischen Bezug, vor allem zur Adhärenz, haben.

Die erste Anforderung aus dieser Gruppe dient dem **Datenschutz** (T1). Ein Ansatz muss die Analyse, Bewertung und Maßnahmendefinition von datenschutzrelevanten Daten implizieren, da mit personenbezogenen Daten gearbeitet werden kann [29]. Die **therapeutische Abdeckung** (T2) verlangt als Anforderung die Betrachtung der individuellen Lebensbedingungen der Zielgruppe, um eine bedarfsgerechte Entwicklung zu ermöglichen und die Umstände entsprechend in die Maßnahmen einzubeziehen [28]. Einher mit der therapeutischen Abdeckung geht die Anforderung der **Patienten Adhärenz** (T3), welche verlangt, konkrete Faktoren für die Messung und Bewertung der Therapietreue zu definieren, um diese explizit mit einbeziehen zu können [19].

Im letzten Abschnitt des Anforderungskatalogs sind diejenigen Anforderungen verortet, die den Rahmen für den Einbezug von präskriptiven Algorithmen stecken.

Hierbei beginnt der Abschnitt mit der Anforderung der **Machbarkeit** (P1) eines Lösungsansatzes. Grundsätzlich sind präskriptive Algorithmen vielseitig und flexibel, was zu einer hohen Entscheidungsfreiheit beim Design dieser führt [30]. Die Lösung sollte jedoch nach dem aktuellen Stand der Technik sowie nach wirtschaftlichen und sicherheitsrelevanten Aspekten umsetzbar sein. (P2) beschreibt die Anforderung der **Nutzenorientierung** des eingesetzten Algorithmus. Vor allem im Hinblick auf die Sicherheit und die Bewertbarkeit dieser eines solchen Algorithmus, sollte der Handlungs- und Einsatzraum eines Algorithmus auf die definierte Zielerreichung spezialisiert werden. Dies reduziert auch den Testraum, in dem die Sicherheit des Patienten / Nutzers gewährleistet werden soll [31]. Da Anforderungen für domänenspezifische Systeme ausschließlich von Domänenexperten kommen können, beschreibt die Anforderung der **Interdisziplinarität** (P3) die zwingende Konsultation von Domänenexperten, was einerseits für die Zielerreichung als auch für die Sicherheit von Nöten ist [8]. Ein weiterer Baustein für die Sicherheit im Prozess ist die Anforderung des **simulativen Testens** (P4) eines Algorithmus. Es muss eine Möglichkeit bestehen, dass der geplante Algorithmus in einer simulativen Umgebung vorab getestet wird, bevor reale Menschen mit ihm konfrontiert werden [32]. Da die Technik, vor allem in Themengebieten wie der künstlichen Intelligenz, rasant voranschreitet, ist (P5) eine Anforderung die **Technologieorientiertheit** und fortwährende Weiterentwicklung des Algorithmus, vor allem im Rahmen der Sicherheit und Genauigkeit. Eine erste Version eines Algorithmus, die den Spezifikationen entspricht, soll nicht die finale Qualität definieren, da täglich neue Erkenntnisse und Ansätze ein besseres und sichereres Erlebnis für die Patienten bedeuten können. Hierbei bedingt (P1) den Handlungskorridor, wie genau (P5) genommen werden soll. (P6) definiert die explizite Betrachtung der **Sicherheit** eines Algorithmus. Für einen Algorithmus müssen überprüfbare Messkriterien erstellt und Alternativmaßnahmen definiert werden, um die Sicherheit während der Nutzung zu gewährleisten [31].

TABLE I. fasst die definierten Anforderungen des Anforderungskatalogs nochmal übersichtlich zusammen.

TABLE I. ANFORDERUNGSKATALOG

Allgemeine Anforderungen	
A1	Iterativ
A2	Flexibel
A3	Nachvollziehbar
A4	messbar
A5	ergebnisorientiert
Therapeutische Anforderungen	
T1	Datenschutz
T2	Therapeutische Abdeckung
T3	Patienten Adhärenz
Anforderungen an präskriptive Algorithmen	
P1	Machbar
P2	Nutzenorientiert
P3	Interdisziplinär
P4	Simulativ
P5	Technologieorientiert
P6	Sicher

C. Bestehende Ansätze

Wie schon bei den Anforderungen wurde zur Identifizierung der bestehenden Ansätze und Methoden eine systematische Literaturanalyse durchgeführt. Hieraus ergaben sich die folgenden Ansätze, die im späteren Verlauf auf die Erfüllung des Anforderungskatalogs hin bewertet wurden. Acht Ansätze wurden dafür identifiziert [33]–[40].

Nach der Auswertung jedes einzelnen Ansatzes wurde eine Übersicht, wie diese in Fig. 2, erstellt, um einen Überblick über den Erfüllungsgrad der einzelnen Anforderungen zu bekommen.

Alle Ansätze erfüllten in der Regel die allgemeinen Anforderungen in mindestens vier von fünf Anforderungen. Dies qualifizierte die gefundenen Ansätze dazu, als Methoden und Vorgehensmodell bezeichnet zu werden, wodurch die Relevanz der Ansätze für die Methodenentwicklung gegeben war.

Im Abschnitt der therapeutischen Anforderungen konnten nur wenige der Ansätze überhaupt die dort definierten Anforderungen ganz oder teilweise erfüllen. Bei der Anforderung (T3), dem Einbezug der Adhärenz, konnten bei allen Ansätzen kein vollständiger und expliziter Einbezug der Adhärenz in den Ansatz identifiziert werden. Dies bildet zudem die größte Lücke im Anforderungskatalog.

Im dritten Abschnitt für präskriptive Algorithmen werden die Anforderungen (P1) und (P2) von allen Ansätzen zumindest teilweise erfüllt. Blick man auf die Anforderungen, die sich implizit und explizit mit der Sicherheit des Algorithmus beschäftigen (P3, P4 und P6), so weisen die Ansätze erhebliche Lücken auf, da bei den meisten identifizierten Ansätzen die Sicherheit als ein inhärenter Bestandteil definiert wurde und keine explizite Betrachtung stattfand.

Zusammenfassend kann festgestellt werden, dass keiner der identifizierten Ansätze den Anforderungskatalog vollständig erfüllen kann. [34] besitzt den höchsten Erfüllungsgrad, besitzt jedoch Schwächen in der Umsetzung und dem Einbezug der therapeutisch relevanten Rahmenbedingungen und der Adhärenz sowie bei der Einhaltung der expliziten Betrachtung der Sicherheit der Lösung.

V. METHODENBESCHREIBUNG

Auf Basis der Erkenntnisse und Anforderungen der des letzten Kapitels, kann eine Methode entwickelt werden, welche alle Anforderungen einbezieht und die definierten Herausforderungen bestmöglich bewältigt.

Zur Entwicklung dieser Methode wurde auf das Method Engineering zurückgegriffen [16], welches die Bestandteile und Inhalte der Bestandteile einer Methode definiert und beschreibt. Bereits in Fig. 3 dargestellt, besteht eine Methode aus mehreren Bausteinen, die untereinander verknüpft sind. So besitzt eine Methode ein Metamodell, welches die Gestaltungsobjekte für die Erstellung eines Modells beschreibt und diese in Beziehung zueinander setzt. Das Vorgehensmodell bildet einen zentralen Baustein einer Methode, da hier die Aktivitäten und Abläufe zwischen den Aktivitäten definiert werden. Für die Ausführung einer Aktivität sind hingegen das Rollen- und Technikmodell von Bedeutung. So beschreibt das Rollenmodell die in der Methode vorhandenen Rollen, welche in den verschiedenen

Aktivitäten miteinander interagieren können. Das Technikmodell hingegen beschreibt die Techniken, die von den Rollen in den Aktivitäten genutzt werden, um die jeweiligen Artefakte innerhalb einer Aktivität zu generieren. Im Ergebnismodell sind die jeweiligen Artefakte definiert.

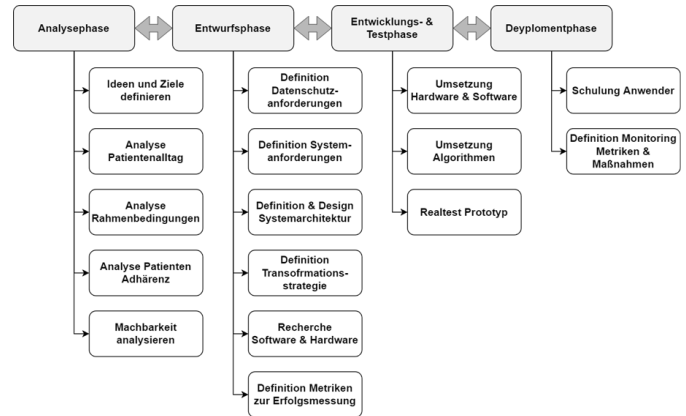


Fig. 6. Funktionsbaum Darstellung des Gesamtvorgehensmodells

In der Fig. 6 ist das Vorgehensmodell der Methode ersichtlich. Es wurde ein Funktionsbaum für die Darstellung hier gewählt, damit das Vorgehensmodell übersichtlich abgebildet werden kann. Mit allen Verzweigungen in Form der BPMN, wäre die Darstellung zu detailliert für den beabsichtigten Zweck geworden.

Das Vorgehensmodell gliedert sich in einen Hauptprozess, welcher den übergeordneten Ablauf und die Rahmenbedingungen definiert und in einen Subprozess (Umsetzung Algorithmen), welcher die Entwicklung des präskriptiven Algorithmus darstellt. Beide Prozesse, sowohl der Haupt- als auch der Subprozess, besitzen die gleichen Phasen, jedoch mit unterschiedlichen Aktivitäten, welche durchlaufen werden müssen. Dazu zählen eine Analyse-, Entwurfs-, Trainings- & Test- sowie die Deploymentphase. Exemplarisch soll auf drei Abschnitte eingegangen werden, die die Herausforderungen der Adhärenz und der Sicherheit der Lösung etwas näher beleuchten.

A. Adhärenz & Datenschutz

Der Einbezug der Adhärenz und des Datenschutzes verteilt sich auf verschiedene Aktivitäten innerhalb des Vorgehensmodells. Patientenalltag und Adhärenz sind zwei grundlegende Aktivitäten der Analysephase, in denen Domänenexperten befragt, Beobachtungen und Befragungen durchgeführt werden und der Alltag exemplarischer Personen der Zielgruppe gezielt mitverfolgt wird. Somit kann sich ein Überblick über die therapeutischen Rahmenbedingungen erstellt werden. Gemessen an den Kriterien der WHO für die Erfassung von Adhärenz [17], [19], wie bereits in Fig. 4 dargestellt, kann ein Gesamtbild über die möglichen Ansätze erstellt werden und eine therapeutische Anforderungsliste erstellt werden. Diese enthält Informationen der WHO Kriterien sowie erste Ideen zu Maßnahmen, wie grundsätzlich eine Unterstützung der Adhärenz aussehen kann.

Somit ist explizit sichergestellt, dass die individuelle Adhärenz und der allgemeine Patientenalltag mit bei der Entwicklung der Lösung einbezogen wird.

Da es sich hierbei im weitreichendsten Fall um medizinische Daten von Personen handelt, sind der Datenschutz und die Medizinproduktbewertung entscheidend für das weitere Vorgehen. Für die Einhaltung des Datenschutzes sind die DSGVO sowie das Bundesdatenschutzgesetz (BDSG) von Bedeutung. Eine Bewertung des Datenschutzes sollte durch einen expliziten Datenschutzbeauftragten geschehen, da dieser die Expertise und Erfahrung hierfür mitbringt. Er kann auch die Risikoanalyse durchführen, die bewertet, welche Risiken wie ausschlaggebend für die schutzbedürftigen Daten der Zielgruppe sind. Gerade im Rahmen der automatisierten Verarbeitung der Daten ist eine Datenschutzfolgeabschätzung nötig, die neben der Identifizierung der Risiken auch Maßnahmen zur Vermeidung oder Reduzierung dieser beinhaltet.

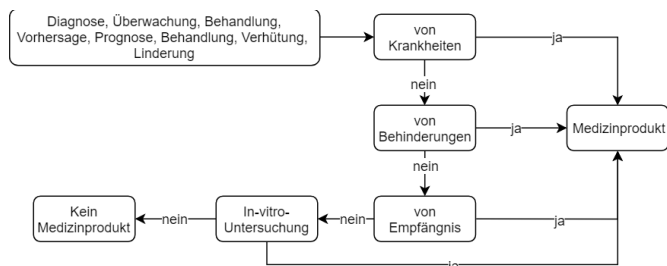


Fig. 7. Übersicht Medizinprodukt Einordnung

Ein weiterer wichtiger Punkt ist die Einordnung und Abgrenzung zu Medizinprodukten [41]. Fig. 7 stellt übersichtlich und prägnant dar, wie die Klassifizierung nach der MDR auszusehen hat. Trifft eine Klassifikation zu und das Produkt landet unter der Rubrik Medizinprodukt, sind weitere Klassifikationen der MDR durchzuführen die aussagen, welche Tätigkeiten für eine Zulassung durchzuführen sind. Im Rahmen dieser Arbeit werden ausschließlich nicht-Medizinproduktenanwendungen betrachtet, da ein Einbezug der Medizinprodukte Regelung ein Prozess wesentlich komplexer gestalten würde und dies keine Anforderung für therapieunterstützende Systeme sein muss und ist.

B. Performance- & Sicherheitsmetriken

Präskriptive Algorithmen können sowohl deterministischer als auch stochastischer Natur sein. In beiden Fällen ist es wichtig, Performance- und Sicherheitsmetriken zu unterscheiden. In der Regel wird die Leistungsfähigkeit eines Algorithmus, bspw. eines neuronalen Netzes, anhand definierter Performancemetriken ermittelt. Dazu können, je nach Anwendungsfall, die mittlere quadratische Abweichung (MSE) bei prädiktiven Regressionen oder der Reward, also ein Qualitätssignal über den aktuellen Zustand, bei präskriptiven Algorithmen dienen. Performance-Kennzahlen sind somit technische Kennzahlen, welche die Zielerreichung des Algorithmus beschreiben. Diese implizieren in der Regel Rahmenbedingungen, wie die Sicherheit des Algorithmus.

Gerade bei präskriptiven Systemen ist die ideale Spezifikation der Zielfunktion (des Rewards, wie er theoretisch erdacht wurde) selten die umgesetzte. Oftmals findet sich eine Design-Spezifikation (die tatsächlich umgesetzte Zielfunktion) vor, die bei der Interaktion des Agenten mit der Umgebung zu einer gänzlich anderen Verhaltensspezifikation (die Zielfunktion, die schließlich vom Agenten versucht wird, zu erreichen) führt. Es

entsteht somit eine Lücke zwischen der Ideal- und der Verhaltensspezifikation, die es mit geeigneten Tests zu schließen bzw. zu mindern gilt. Denn allein die Performance-Metriken können ein Erreichen eines Optimums zwar beschreiben, jedoch nicht dessen Grad der Genauigkeit zur Idealspezifikation widerspiegeln [42].

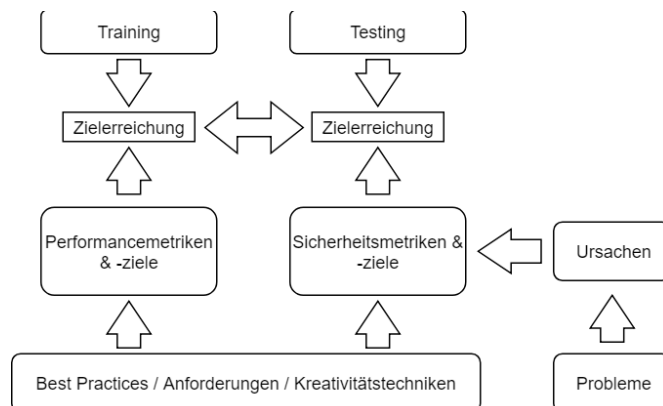


Fig. 8. Trennung zwischen Performance- und Sicherheitsmetriken

Deshalb wurden im Design der Methode Sicherheitsmetriken zur Nutzung von präskriptiven Algorithmen definiert, die potentielle Probleme und deren Ursachen mit einbeziehen (siehe Fig. 8), um Testszenarien bzw. -verfahren zu definieren, die zusätzlich zu den Performance-Metriken herangezogen werden. Sie sollen somit sicherheitskritisches Verhalten des Agenten in der Umgebung reduzieren.

TABLE II. SICHERHEITSRISIKEN BEI PRÄSKRIPTIVEN ALGORITHMEN [8], [31]

Risiko	Beschreibung
Safe Interruptibility	Ein System muss unterbrechbar sein.
Avoiding Side Effects	Es dürfen bei der Erreichung des Ziels keine negativen Effekte auftreten.
Absent Supervisor	Das Verhalten in der realen Welt soll dem Verhalten in der Testwelt entsprechen.
Reward Hacking	Die Zielfunktion darf nicht vom System beeinflusst werden können.
Self-Modification	Veränderung des Verhaltens darf Sicherheit und Zielerreichung nicht beeinflussen.
Distributional Shift	Veränderungen der Umgebung dürfen zu keinen Sicherheitsrisiken führen.
Robustness to Adversaries	Auch in Multi-Agenten-Systemen soll die Zielfunktion aller gewahrt bleiben, auch wenn die Ziele gegensätzlich sind.
Safe Exploration	Ein System soll sich in unbekannten Situationen so verhalten, dass kein Schaden entsteht und Risiken minimiert werden.

Scalable Oversight	Es soll sein Ziel versuchen zu erreichen, auch wenn es dies nicht überprüfen kann.
--------------------	--

Durch die gestiegene Popularität von präskriptiven Algorithmen, wie bspw. dem bestärkenden Lernen, durch z. B. spielende Agenten wie OpenAI Five [43] oder AlphaStar [44], wurde vermehrt nach Sicherheitsrisiken dieser Ansätze geforscht. Es ergaben sich neun hauptsächliche Risiken, welche in der TABLE II. kurz vorgestellt werden. All diese Risiken müssen bei der Definition der Sicherheitsmetriken mit in betracht gezogen werden, wenn auch nicht alle bei jedem System eine Rolle spielen.

Beispielhaft muss bei einem *distributional shift*, also bei einer Änderung der Umgebung, z. B. beim Übergang von einer simulativen in eine reale Welt, der Agent trotzdem sichere Entscheidungen treffen, auch wenn er die aktuelle Umgebung nicht kennt.

So ist es in einer Trainingswelt möglich, den Weg gerade zu beschreiten, um sein Ziel zu erreichen. In der realen Welt ist auf diesem Weg ein Hindernis, z. B. eine Baustelle. Auch wenn dieser Fall niemals während des Trainings aufgetreten ist, muss das System in der Lage sein, auf solche unvorhergesehenen Situationen zu reagieren.

Auch wenn nicht alle Probleme und Sicherheitsrisiken mit Testfällen nachgestellt werden können, besteht bei lernenden Systemen immer die Möglichkeit, dass diese sich generalisieren, also Situationen erkennen, obwohl sie diese nie sahen.

C. Algorithmtests

Bei der Durchführung der Testphase für die Algorithmen müssen verschiedenste Sicherheitsrisiken (wie aus TABLE II. ersichtlich) ausgeschlossen oder abgedeckt werden. Dabei sind die in der TABLE II. vorhandenen Sicherheitsrisiken lediglich generalisiert und abstrakt, weshalb ein Domänenbezug zur konkreten Erstellung von Testfällen erforderlich ist.

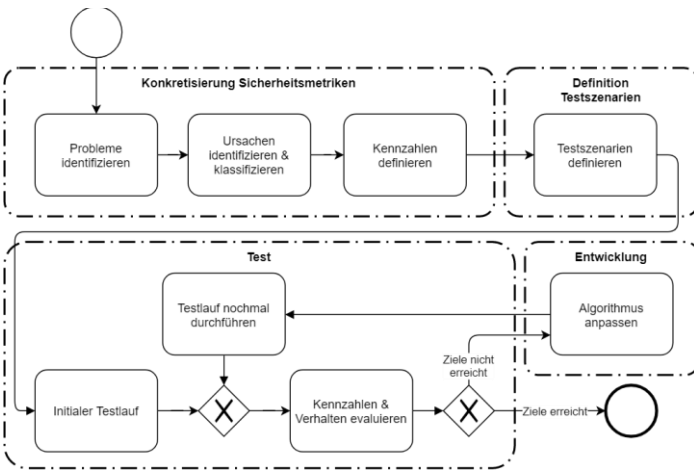


Fig. 9. Definition von Sicherheitsmetriken und Ablauf von Testläufen

Fig. 9 stellt den Ablauf für die Erstellung und Durchführung von Testszenarien vor. In der Abbildung ist zusätzlich ersichtlich, dass die verschiedenen Gruppen der Abbildung in verschiedenen Aktivitäten im Vorgehensmodell untergebracht sind und

somit einen integrierten Prozess definieren. Aus den zuvor erstellten Kennzahlen werden Testszenarien erstellt, die mögliche Sonderfälle in der realen Welt versuchen abzudecken. Dabei werden nach jedem Testdurchlauf die Kennzahlen und ggf. weitere Analysen, wie Verhaltensanalysen mit t-sne [45] oder formale Verifikationen [46], durchgeführt, um die Handlungen besser zu verstehen und Anpassungen des Algorithmus, z. B. mit sicheren Explorationen [47] durchzuführen.

Schließt das System die Testfälle erfolgreich ab, liegen die Kennzahlen im akzeptierten Bereich und finden sich keine auffälligen Verhaltensmuster mehr, kann der Agent als getestet gelten.

VI. EVALUATION

Die Evaluation der entwickelten Methode fand im Rahmen eines Forschungsprojektes statt und wurde mit den darin beschäftigten Domänenexperten aus der Pflege durchgeführt und evaluiert. Dieses hatte zum Ziel, seniorengerechte Stadtquartiere zu entwickeln. Ein Element des Forschungsprojektes ist die Entwicklung einer adaptiven Navigation, welche Menschen mit 65 Jahren oder älter, wohnhaft in einer Pflegeeinrichtung, ermöglicht, auf Basis ihrer Kompetenzen (Einschränkungen), wie bspw. Rollatoren, Rollstühle oder andere Gehbehinderungen, den für sie am besten passenden Weg zu bestreiten, um an ihr Ziel zu kommen. Dies ist ein Ansatz, der sowohl dem Projektziel dient als auch eine Aktivitätsunterstützung nach sich zieht, was zu einer Bewegungsmotivation führt. Diese Bewegungsmotivation kann sich positiv auf die Adhärenz bei chronischen Erkrankungen auswirken. Somit war der Einsatz der Methode passend und für die Entwicklung dieses adaptiven Routings zielführend.

Neben der Adhärenz konnte vor allem die Sicherheit eines solchen adaptiven Systems für den Einsatz in der realen Welt erhöht werden. Dies referenziert auf die *RQ3*, die Optimierungen durch das entwickelte System fordert. Da beim Routing vor allem die Bodenbeschaffenheiten und Hindernisse eine Rolle spielten, war es besonders wichtig, dass die Senioren vor allem Wege vorgeschlagen bekamen, die ihrer Kompetenz nach am besten begehbar waren. So sind Waldwege oder sandhaltige Untergründe vor allem für Rollstuhlfahrer besonders problematisch und anstrengend, wohingegen Personen mit Gehstöcken oder anderen nicht sitzenden Kompetenzen, ab und an Sitzmöglichkeiten, selbst bei kürzeren Strecken, benötigen.

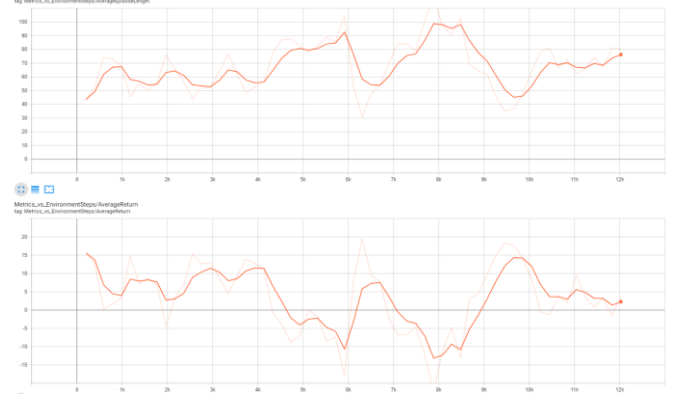


Fig. 10. Wegstrecke (oben) und Reward (unten) eines Testlaufs; beide pendelten sich im Laufe der Trainingsphasen auf ein Optimum ein

Wie aus Fig. 10 ersichtlich, geben der Reward und die Wegstrecke als Performance Indikatoren keinen Aufschluss über die Sicherheit des Systems. Lediglich das Erreichen eines gewissen Optimums kann nachvollzogen werden. Erst unter Einbezug einer Verhaltensanalyse kann das eigentliche Verhalten konkret nachvollzogen werden und ggf. Probleme aufgedeckt werden.

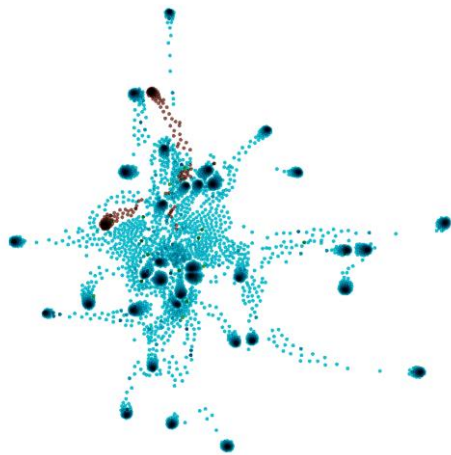


Fig. 11. Verhaltensanalyse mit t-sne (in Anlehnung an [48])

Wirft man jedoch einen Blick auf Fig. 11, können zwei Verhaltensausreißer nachvollzogen werden. Bestimmte Ballungen in der Abbildung entsprechen Zielen (Apotheken, Märkte etc.), welche erreicht werden sollten, die einzelnen Punkte sind Zustände, in denen der Agent sich aufhielt. Türkis hinterlegte Punkte sind Zustände, in denen eine Person mit der Kompetenz Rollstuhl auf einem befestigten, asphaltierten Weg unterwegs war. Die zwei braunen Ballungen hingegen sind Zustände, in denen die Testperson auf sandigen Waldboden unterwegs war. Reward und Streckenmessungen konnten diese offensichtlich nicht optimalen Wege nicht kenntlich machen. Mit der Wahl von richtigen Metriken und Metadaten war es jedoch in der Verhaltensanalyse möglich einen Einblick in die Entscheidungen des Agenten zu erhalten.

Nach der Erkenntnis wurden die entsprechenden Zustände rekonstruiert, nachvollzogen und der Agent entsprechend angepasst, damit solche Situationen nicht mehr auftreten. Final waren die Zustände nicht mehr aufgetreten. Was hier eine relativ gut kontrollierbare Umgebung ist, kann bei anderen Anwendungsfällen wesentlich komplexer werden (mehr Zustände, mehr Metadaten, mehr Entscheidungsmöglichkeiten). Ein Einblick in das Verhalten von Agenten wurde schon häufiger in der Vergangenheit genutzt [48] und verhalf dem Verständnis und der Sicherheit von Systemen beizutragen. Dabei ist die Verhaltensanalyse mit t-sne nur eine Möglichkeit, Einblick in Systeme zu erhalten. Distinctly Selective Neurons [48] sind bspw. eine weitere unter vielen Möglichkeiten, bestimmtes Verhalten zu verstehen und zu analysieren.

VII. RESUMÉ & AUSBLICK

In diesem Beitrag wurde eine Methode entwickelt, welche sowohl die Adhärenz bei chronischen Patienten als auch die Sicherheit explizit mit einbeziehen sollte, um

therapieunterstützende präskriptive Systeme bauen zu können, deren Fokus auf der Sicherheit und Unterstützung im Alltag von Patienten liegt.

Um dieses Ziel zu erreichen, musste eine systematische Literaturanalyse durchgeführt werden, um die Herausforderungen und bestehenden Ansätze sowie die benötigten Anforderungen für einen solchen Ansatz zu identifizieren. Daraufhin wurde ein Anforderungskatalog erstellt und die bestehenden Ansätze dagegen geprüft. Heraus kam, dass keiner der Ansätze den Anforderungskatalog vollständig ausfüllt.

Um die Entwicklung eines präskriptiven Therapieunterstützungssystems systematisch durchführen zu können, wurde eine Methode gebaut, die nach dem Vorgehen des Method Engineering erstellt wurde. Diese beinhaltet ein Metamodell, ein Vorgehensmodell, ein Technik- und ein Rollenmodell sowie ein Ergebnismodell.

Die Evaluation des Ansatzes in einem Forschungsprojekt aus dem Pflegebereich, zusammen mit den Experten aus der Pflege, bot die Möglichkeit, Optimierungen der dort entwickelten Ansätze zu erkennen und durch den Einsatz der Methode umzusetzen. Dabei konnte mit dem Fokus auf die Sicherheit der Entscheidungsunterstützung eine therapieunterstützende Navigation für Senioren sicher umgesetzt werden.

Zukünftig sollen zusätzliche Evaluationen im Pflegebereich mit anderen Anwendungsfällen und Domänenexperten durchgeführt werden, die eine Generalisierung der Methode nachweisen zu können.

Die Domänen- und Anwendungsfallunabhängigkeit des Ansatzes lässt zukünftig weitere Anwendungsmöglichkeiten zu, die weitaus komplexere Zusammenhänge abbilden können.

REFERENCES

- [1] Allied Market Research (Hg.), "mHealth (mobile health) industry market size projection from 2012 to 2020 (in billion U.S. dollars)", Statista, 2013. Online: <https://www.statista.com/statistics/295771/mhealth-global-market-size/>, zuletzt geprüft am 02.08.2019.
- [2] Berger, Roland, "Global digital health market from 2015 to 2020, by major segment (in billion U.S. dollars)", Statista, 2016. Online: <https://www.statista.com/statistics/387867/value-of-worldwide-digital-health-market-forecast-by-segment/>, zuletzt geprüft am 02.08.2019.
- [3] Initiative Neue Qualität der Arbeit (BAuA) (Hg.), "Digitalisierung in der Pflege. Wie intelligente Technologien die Arbeit professionell Pflege in der verändern", 2018.
- [4] Rothgang, Heinz; Müller, Rolf; Unger, Rainer, "Themenreport „Pflege 2030“. Themenreport „Pflege 2030“, Bertelsmann Stiftung, 2012.
- [5] DAK (Hg.), "Umfrage zur Überforderung durch eine Pflege Tätigkeit in Deutschland nach Altersgruppen im Jahr 2015", Statista, 2015. Online: <https://de.statista.com/statistik/daten/studie/481570/umfrage/ueberforderung-durch-eine-pflegetaetigkeit-in-deutschland-nach-altersgruppen/>, zuletzt geprüft am 02.08.2019.
- [6] Hehner, Steffen; Biesdorf, Stefan; Möller, Manuel, "Digitalisierung im Gesundheitswesen: die Chancen für Deutschland", McKinsey & Company, 2018.
- [7] Hassan, Qusay F.; Khan, Atta ur Rehman; Madani, Sajjad A., "Internet of things. Challenges, advances, and applications", Boca Raton: Taylor & Francis, CRC Press (Chapman & Hall/CRC computer & information science series), 2018.
- [8] Brundage, Miles; Avin, Shahar; Clark, Jack; Toner, Helen; Eckersley, Peter; Garfinkel, Ben et al., "The Malicious Use of Artificial Intelligence. Forecasting, Prevention, and Mitigation", 2018. Online: <https://arxiv.org/pdf/1802.07228.pdf>.

- [9] World Health Organization, "Classification of Digital Health Interventions v1.0", 2018.
- [10] Österle, Hubert; Winter, Robert; Brenner, Walter (Hg.), "Gestaltungsorientierte Wirtschaftsinformatik. Ein Plädoyer für Rigor und Relevanz," Nürnberg, Infowerk, 2010.
- [11] Österle, Hubert; Becker, Jörg; Frank, Ulrich; Hess, Thomas; Karagiannis, Dimitris; Krcmar, Helmut et al., "Memorandum zur gestaltungsorientierten Wirtschaftsinformatik," Schmalenbachs Z betriebswirtsch Forsch, vol. 62, pp. 664–672, 2010. DOI: 10.1007/BF03372838.
- [12] Vom Brocke, Jan; Simons, Alexander; Niehaves, Björn; Riemer, Kai; Plattfaut, Ralf; Cleven, Anne, "Reconstructing the giant: On the importance of rigour in documenting the literature search process," ECIS, vol. 161, pp. 2206–2217, 2009.
- [13] Aronson, Elliot; Wilson, Timothy D.; Akert, Robin M., "Sozialpsychologie," vol. 6, München, Pearson Studium (ps Psychologie), 2011.
- [14] Schnell, Rainer; Hill, Paul B.; Esser, Elke, "Methoden der empirischen Sozialforschung," vol. 5, München, 1995.
- [15] Rupp, Chris; Simon, Matthias; Hocker, Florian, "Requirements Engineering and Management," HMD, vol. 46 (3), pp. 94–103. 2009. DOI: 10.1007/BF03340367.
- [16] Gutzwiller T.A., "Das CC RIM-Referenzmodell für den Entwurf von betrieblichen, transaktionsorientierten Informationssystemen," Heidelberg, Physica-Verlag HD, vol. 54, 1994.
- [17] Schäfer C., "Grundlagen der Patientencompliance und Adhärenz," Patientencompliance, Wiesbaden, Springer Fachmedien Wiesbaden, pp. 13–40, 2017.
- [18] Grønhaug G., "Addressing the elephant in the room: a possible new way to increase patient adherence to medical advice," Patient preference and adherence, pp. 1083–1089, 2017. DOI: 10.2147/PPA.S138716.
- [19] Weltgesundheitsorganisation (2003): Adherence to long-term therapies. Evidence for action. Geneva.
- [20] Brown M.T., Bussell J.K., "Medication adherence: WHO cares?," Mayo Clinic proceedings, vol. 4, pp. 304–314, 2011. DOI: 10.4065/mcp.2010.0575.
- [21] Prinz, "Wie erfülle ich die regulatorischen Anforderungen für Künstliche Intelligenz Anwendungen in der Medizin?," Verband deutscher Elektrotechnik Elektronik Informationstechnik e.V., Frankfurt, 2019.
- [22] Wenner, "Risk Management for Medical Devices," 2019. Available from: <https://meso.vde.com/risk-management-medical-devices/>, 09.19.
- [23] Puget, "Analytics Landscape," 2015. Available from: https://www.ibm.com/developerworks/community/blogs/jfp/entry/Analytics_Models?lang=en, 08.18.
- [24] Mauerer, "Was ist was bei Predictive Analytics?," 2018. Available from: <https://www.cio.de/a/was-ist-was-bei-predictive-analytics,3098583,4>, 08.18.
- [25] Leike J., Krueger D., Everitt T., Martic M., Maini V., Legg S., "Scalable agent alignment via reward modeling: a research direction," ArXiv, abs/1811.07871, 2018.
- [26] Schreiegg/Prinz, "Software Classification: Insights in EU Guidance," 2019. Available from: <https://meso.vde.com/software-classification-guidance/>, 09.19.
- [27] Bourque P., Fairley R.E., "Guide to the software engineering body of knowledge," vol. 3, Los Alamitos, Calif., IEEE Computer Society, 2014.
- [28] Riedl, Rainer, "Strategische Planung von Informationssystemen. Methode zur Entwicklung von langfristigen Konzepten für die Informationsverarbeitung," Wirtschaftswissenschaftliche Beiträge, vol. 51, Heidelberg: Physica-Verlag HD, 1991.
- [29] Bundesministerium der Justiz und für Verbraucherschutz (2018): Bundesdatenschutzgesetz. BDSG, 2018.
- [30] Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron, "Deep learning," 2nd ed., Cambridge, Massachusetts, London, England, MIT Press, 2016.
- [31] Amodei D., Olah C., Steinhardt J., Christiano P., Schulman J., Mané D., "Concrete Problems in AI Safety," ArXiv, abs/1606.06565, 2016.
- [32] Sutton R.S., Barto A.G., "Reinforcement learning," 2nd ed., Cambridge, Mass., MIT Press, 2018.
- [33] Zhang L., Alharbe N.R., Atkins A.S., "A self-adaptive distributed decision support model for Internet of Things applications," Transactions of the Institute of Measurement and Control, vol. 4, pp. 404–419, 2017.
- [34] Dziak D., Jachimczyk B., Kulesza W., "IoT-Based Information System for Healthcare Application: Design Methodology Approach," Applied Sciences, vol. 6, pp. 596, 2017.
- [35] Yoon J., Davtyan C., van der Schaar M., "Discovery and Clinical Decision Support for Personalized Healthcare," IEEE journal of biomedical and health informatics, vol. 4, pp. 1133–1145, 2017.
- [36] Forkan A.R.M., Khalil I., "A clinical decision-making mechanism for context-aware and patient-specific remote monitoring systems using the correlations of multiple vital signs," Computer methods and programs in biomedicine, pp. 1–16, 2017.
- [37] EMC Education Services, "Data science & big data analytics. Discovering, analyzing, visualizing and presenting data," Indianapolis, IN: John Wiley & Sons, 2015.
- [38] Fortino G., Russo W., "ELDAMeth: An agent-oriented methodology for simulation-based prototyping of distributed agent systems," Information and Software Technology, vol. 6, pp. 608–624, 2012.
- [39] Foglyano K.M., Schnellenberger J.R., Kobetic R., Lombardo L., Pinault G., Selkirk S., Makowski N.S., Triolo R.J., "Accelerometer-based step initiation control for gait-assist neuroprostheses," Journal of rehabilitation research and development, vol. 6, pp. 919–932, 2016.
- [40] Chapman et al., "CRISP-DM 1.0," 2000. Available from: <https://www.the-modeling-agency.com/crisp-dm.pdf>, 04.19.
- [41] European Union (EU), "2017/746 Medical Device Regulation. MDR," 2017. Available from: <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32017R0745&from=EN>, 10.18.
- [42] Ortega/Maini, "Building safe artificial intelligence: specification, robustness, and assurance," 2018. Available from: <https://medium.com/@deepmindsafetyresearch/building-safe-artificial-intelligence-52f5f75058f1>, 06.19.
- [43] OpenAI Inc., "OpenAI Five," 2018. Available from: <https://openai.com/blog/openai-five/>, 09.19.
- [44] Vinyals et al., "AlphaStar: Mastering the Real-Time Strategy Game StarCraft II," 2019.
- [45] Zahavy T., Zrihem N.B., Mannor S., "Graying the black box: Understanding DQNs," ArXiv, abs/1602.02658, 2017.
- [46] Qin, C., Dvijotham, K., ODonoghue, B., Bunel, R., Stanforth, R., Gwal, S., Uesato, J., Swirszcz, G., Kohli, P., "VERIFICATION OF NON-LINEAR SPECIFICATIONS FOR NEURAL NETWORKS," International Conference on Machine Learning, vol. 97, Long Beach, California, USA, 2019.
- [47] Eysenbach B., Gu S., Ibarz J., Levine S., "Leave no Trace: Learning to Reset for Safe and Autonomous Reinforcement Learning," ArXiv, abs/1711.06782, 2017.
- [48] Jaderberg M. et al., "Human-level performance in first-person multiplayer games with population-based deep reinforcement learning," ArXiv, abs/1807.01281, 2018.